



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>G06F 19/00</b>	<b>A1</b>	(11) International Publication Number: <b>WO 00/51058</b>
		(43) International Publication Date: 31 August 2000 (31.08.00)

(21) International Application Number: PCT/US00/04926

(22) International Filing Date: 25 February 2000 (25.02.00)

(30) Priority Data:  
09/259,065 26 February 1999 (26.02.99) US

(71) Applicant: GENERAL SCANNING, INC. [US/US]; 500 Arsenal Street, Watertown, MA 02472 (US).

(72) Inventors: ZELENY, Rolland; 34 Magnolia Road, Sharon, MA 02067 (US). SCHERMER, Mack; 284 Lake Street, Belmont, MA 02478 (US). ACHIN, Ron; 23 Houde Street, Marlborough, MA 01752-1056 (US).

(74) Agents: CESARI, Robert, A. et al.; Cesari and McKenna, LLP, 30 Rows Wharf, Boston, MA 02210 (US).

(81) Designated States: JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

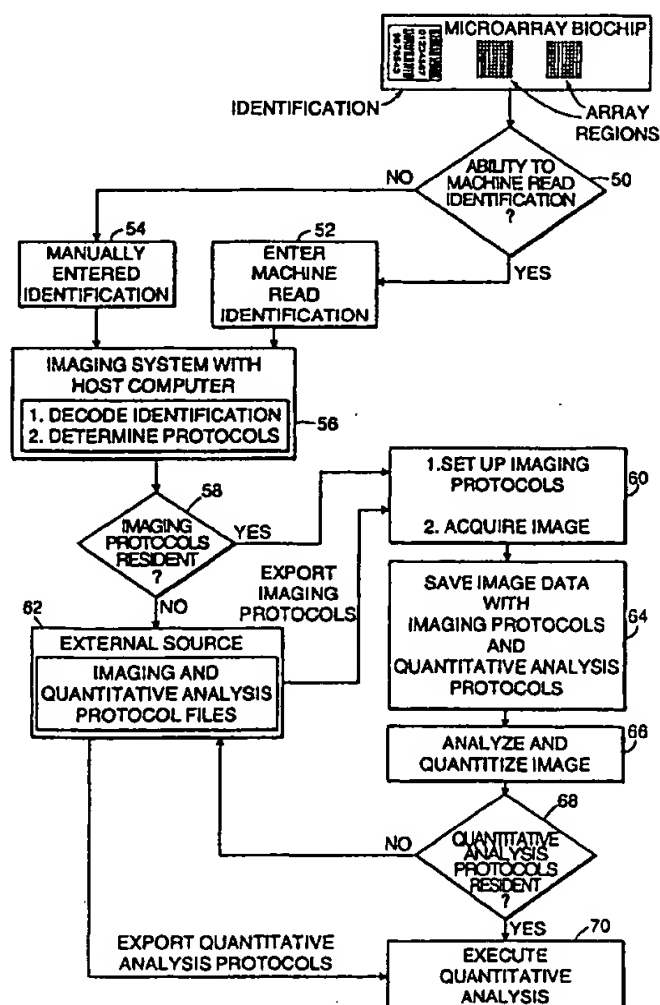
## Published

*With international search report.**Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: AUTOMATIC IMAGING AND ANALYSIS OF MICROARRAY BIOCHIPS

## (57) Abstract

A system for scanning biochip arrays includes a unique image array identifier recorded for each array, and a computer-stored record corresponding to each identifier and containing the parameters of the experiment in the array identified by the identifier. The system further includes means for accessing a protocol library to retrieve the scanning protocols associated with the identified arrays and then scanning the arrays in accordance with the respective protocols. The resulting image maps generated by the scanners are stored in locations corresponding to the associated identifiers.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

# **AUTOMATIC IMAGING AND ANALYSIS OF MICROARRAY BIOCHIPS**

## **FIELD OF THE INVENTION**

This invention relates to the imaging and analysis of experiments performed on  
5 microarray biochips. More particularly it relates to the automatic control of microarray  
scanners and selection of the proper protocols for analysis of the image maps provided  
by the scanners.

## **BACKGROUND OF THE INVENTION**

Microarray biochips are being increasingly used for the performance of large  
10 numbers of closely related chemical tests. For example, to ascertain the genetic differ-  
ences between lung tumors and normal lung tissue one might deposit small samples of  
different cDNA sequences under a microscope slide and chemically bond them to the  
glass. ten thousand or more such samples can easily be arrayed as dots on a single mi-  
croscope slide using mechanical microarraying techniques. Next, sample mRNA is ex-  
15 tracted from normal lung tissue and from a lung tumor. The mRNA represents all of  
the genes expressed in these tissues and the differences in the expression of mRNA  
between the diseased tissue and the normal tissue can provide insights into the cause of  
the cancer and perhaps point to possible therapeutic agents as well. The  
"probe" samples from the two tissues are labeled with different fluorescent dyes. A  
20 predetermined amount of each of the two samples is then deposited on each of the mi-  
croarray dots where they competitively react with the cDNA molecules. The mRNA  
molecules that correspond to the cDNA strands in the array dots bind to the strands and  
those that do not are washed away.

The slide is subsequently processed in a scanner that illuminates each of the  
25 dots with laser beams whose wavelengths correspond to the fluorescences of the label-

- 2 -

ing dyes. The fluorescent emissions are sensed and their intensity measured to ascertain for each of the array dots the degree to which the mRNA samples correspond to the respective cDNA sequences. In the experiment outlined above the image scanner separately senses the two fluorescences and thereby provides separate maps of the reactions  
5 of the mRNA extracted from the normal and tumorous tissues. The scanner generates an image map of the array, one for each of the fluorescences. The maps are ultimately analyzed to provide meaningful information to the experimenter.

Microarray biochips are available in a variety of form factors and they may contain one or more different fluorescence labels. The reagents involved in the chemical reactions in the array dots are typically biological samples such as DNA, RNA,  
10 peptides, proteins or other organic molecules. The biochips might be used for diagnostics, screening assays, genetics and molecular biology research. They may include, in addition to the test dots, calibration dots containing known amounts of the fluorescent materials. Scanning of the latter dots thus serves to calibrate the readings obtained  
15 from the test dots.

In order to obtain accurate information from the scanning of a biochip array, it is important to know which fluorescence materials have been used, in where the array is located, and the locations of the calibration dots, if any. It is important to know the fluorescent materials in order to use the correct wavelengths in illuminating the dots  
20 and to filter the correct wavelengths of the fluorescent omissions. Furthermore, it is advantageous to excite the fluorescences with a high intensity so as to provide the maximum signal to the fluorescence detector. However, the intensity must be kept below the level at which the fluorescence is saturated or the fluorescent material is degraded and this depends on the particular fluorescent material that is used.

25 Furthermore, analysis of raw data collected by the scanner must be performed in accordance with protocols that may vary in accordance with experiment parameters. Prior to the present invention entry of the scanning and analysis protocol has been performed manually. This involves significant operator time and, further, is a source of errors in the scanning and analysis procedure.

## SUMMARY OF THE INVENTION

In accordance with the invention an identifier corresponding to each experiment is imprinted on the biochip. The identifier is machine readable and, preferably, also human readable. For example, the identifier may be a number with numerals imprinted on the slide along with a bar code representation of the number. The experiment identifier is imprinted on the chip prior to the deposition of the array experiment. A file folder (i.e., "directory") is opened in a computer system and is logically linked to the array identifier. An operator may enter into that folder the various parameters of the experiment array, e.g., a map of the reagents deposited in the array, identification of the fluorescent tags and the reagents to which they are bonded, and also the locations of any calibration dots on the chip. The operator may also enter into the folder an identification of the scanning and processing protocols to be used in connection with the scanning process. Preferably, however, the system is programmed to retrieve the information from the biochip. The protocols are therefore identified, or even described, in the machine-readable code on the biochip.

Accordingly, when the biochip is subsequently installed in a scanner, the scanner can scan the bar code version of the identifier. The system can then automatically open the requisite file folders and obtain the scanning and analysis protocols. The scanner is then set up to operate in accordance with the retrieved protocol. The resulting image map of the scanned data is stored in one or more files in the chip's file folder. The analysis protocol is subsequently retrieved and the image map processed in accordance with the latter protocol provide an array of output data.

If the system cannot locate an appropriate protocol, either for scanning or processing, it prompts the operator to identify an appropriate stored protocol or to enter a new protocol, if required.

With the foregoing arrangement the requirement for operator intervention is significantly reduced. This speeds up the overall process and, perhaps more important, reduces the probability of error that would be encountered error due to improper entry of scanning and processing protocols.

## BRIEF DESCRIPTION OF THE DRAWINGS

The invention description below refers to the accompanying drawings, of which:

Fig. 1 depicts a biochip containing two experiment

5 Fig. 2 is an enlargement of a portion of Fig 1.

Fig. 3 is a diagram of a system operating in accordance with the invention; and

Fig. 4 is a flow diagram of the process of the invention.

## DETAILED DESCRIPTION OF AN ILLUSTRATIVE EMBODIMENT

10

In Fig. 1, a microarray biochip 10 is embodied in the form of a microscope slide than having two array regions 12 and 14. Each of the array regions involves an experiment comprising an array of tests each of which involves a different combination of reagents. The tests are performed by depositing in drops of reagents in locations depicted as dots. The regions 12 and 14 comprise different experiments, in which case the  
15 tests in region 12 will be different from those in region 14.

The biochip 10 has also been imprinted with experiment identifiers 16 and 18 relating to the experiments performed in the regions 12 and 14. Specifically, each of the identifiers comprises a numeral (16a, 18a) and a corresponding bar code (16b, 17b)  
20 representation of the numeral, as illustrated in the drawing. The identifier may be a number in which some of the digits identify experiment parameters of the array, others identify the source of the arrays, others may identify the scanning and analysis protocols or even provide the operating parameters for those protocols, and still others identify the array itself.

25 A typical array region is depicted in greater detail in Fig. 2. The region comprises an array of test spots 20, and optionally, a series of control spots 22 and focus targets 24. The control spots 22 and the targets 24 are used at the beginning of the scanning process to calibrate and/or adjust the scanner.

## CLAIMS

- 1 1. A system for processing biochip arrays each of which comprise an experiment  
2 consisting of a matrix of tests performed on a biochip, the system comprising  
3 A. a plurality of biochips, each biochip carrying a one or more ar-  
4 ray identifiers each of which uniquely identifies an array on the  
5 bio chip,  
6 B. means for recording, in a computer file corresponding to each  
7 identifier, the parameters of the experiment in the array identified  
8 by the identifier,  
9 C. a library stored in a computer and containing protocols for  
10 scanning the arrays of various experiments,  
11 D. a scanner for scanning biochip arrays and including means for  
12 retrieving from biochip the identifiers recorded thereon,  
13 E. means responsive to the retrieved identifiers for selecting from  
14 the library scanning protocols for the identified arrays,  
15 F. means for controlling the operation of the scanner in accordance  
16 with the retrieved imaging protocols, and  
17 G. means for recording in a computer, in a location corresponding  
18 to each identifier, an image map generated by the scanner in  
19 scanning the corresponding experiment array.
- 1 2. The system defined in claim 1, including an analysis for analyzing the recorded  
2 image maps, said analyzer including means for selecting from the library the analysis  
3 protocols corresponding to the array identifiers corresponding to the image map.
- 1 3. A method of processing biochip arrays, each array comprising one or more ex-  
2 periments, each of which consists of matrix of tests deposited on the biochip, said  
3 method comprising the steps of:  
4 A. imprinting on each biochip a machine-readable identifier corresponding  
5 to each of said arrays,  
6 B. recording in a computer system a file containing, for each identifier, the

- 5 -

As shown in Fig. 3, a system embodying the invention, includes an identifier printer 30 that imprints the experiment identifiers on the microchips. The chips are loaded into an arrayer 32, where the test spots are deposited to form the experiment arrays.

5 The biochip subsequently passes to a scanner 42 which first scans the bar codes 16b and 18b. The system 34 responds by opening file folders 36 that are logically linked to the identifiers. In the folders 36 it opens files containing the information in the identifiers. If the bar codes identify protocols contained in a protocol library 35, or if the protocols are contained in the identifier, the system copies the protocols into files in  
10 the biochip's corresponding file folders. The system then sets up the scanner to operate in accordance with the retrieved scanning protocols.

The image maps obtained by the scanner are stored in image maps 44 contained in the folders 36. Finally, an analyzer 45 in the host computer system analyzes the contents of the image maps 44 to provide output tables 46 also stored in the folder 36.  
15 In generating an output table the analyzer 45 uses the analysis protocols contained in folders 36.

A flow diagram of the scanning and image analysis operations is depicted in Fig 4. When a biochip 10 is loaded into the scanner 42, the scanner, in step 50, attempts to read the array identifiers on the chip. If it can read the identifiers, it transmits them to  
20 the host computer system 34 as shown at step 52. If the scanner cannot read the identifiers, it prompts the operator to enter them manually as indicated in step 54. At step 56 the host computer decodes the identifiers and ascertains the protocols that are to be used. Assuming the identifier points to protocols, the computer system, at step 58 ascertains whether the protocol library 38 contains a protocol identified in step 56. If it  
25 does, the protocol is applied to the scanner 42, which, at step 60, sets up the imaging protocols and acquires the microarray image map. If the library 38 does not contain the identified protocol, it proceeds to step 62 to obtain the protocols from an external source (not shown). The retrieved imaging protocol is then passed to the scanner 42, which operates as indicated in step 60. In step 64, the image maps are recorded in an  
30 image file 44.



- 6 -

At a subsequent time, the analysis program is invoked at step 66 and at step 68 the system determines whether the appropriate analysis protocol is resident in the protocol library 38. If it is, the system executes the quantitative analysis at step 70. If the library 38 does not contain the appropriate protocol, the external source is examined,  
5 and, it passes the protocol to step 70.

It will be apparent that the foregoing system operates with minimal operator intervention and, therefore, it processes the biochips materially faster than prior arrangements. Moreover, by tying each biochip array to the appropriate protocols at an early stage of the process, it reduces the errors that would otherwise be encountered during  
10 the imaging and analysis stages of the procedure.

It should be understood that the invention is applicable to biochips other than the biochip illustrated herein. For example, it may be used with "wet" biochips, which contain reagents in liquid form in arrayed wells. Further, as used herein the term "biochip" also applies to materials that are not deposited in ordered arrays, e.g., tissues af-  
15 fixed to slides, which are scanned for fluorescent or other labels.

What is claimed is:

- 8 -

- 7 parameters of the experiments in the arrays identified by the array  
8 identifiers on the biochip.
- 9 C. entering into a library in said computer system, the protocols for  
10 for scanning the biochip rays for various experiments,  
11 D. in response to each array identifier, selecting from said library the  
12 imaging protocols for the experiment in the array identified  
13 by the identifier, and  
14 E. controlling a scanner in accordance with the selected scanning  
15 protocol to scan the identified array and record in said computer system.  
16 A corresponding image map of the array.
- 1 4. The method defined in claim 3, including the steps of  
2 A. selecting from the protocol library a protocol for the analysis of each of  
3 said image maps, and  
4 B. Analyzing the image map in accordance with each selected protocol  
5 analyzing the image map.

1/3

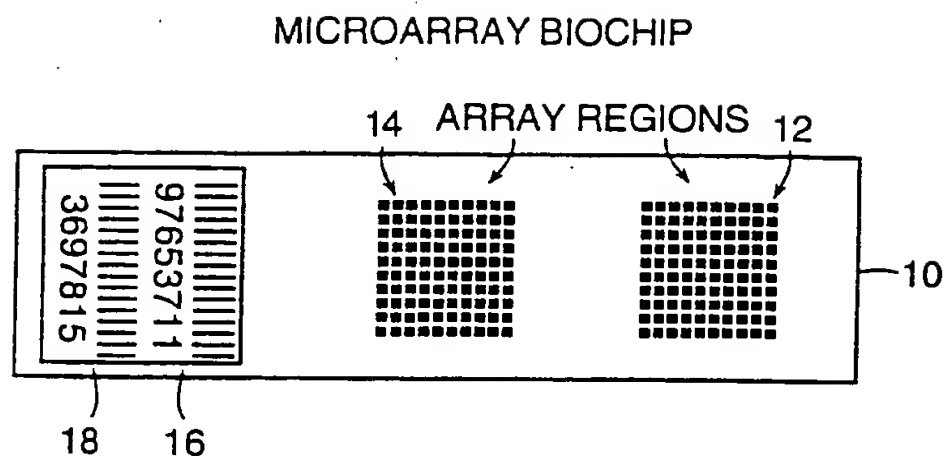


FIG. 1

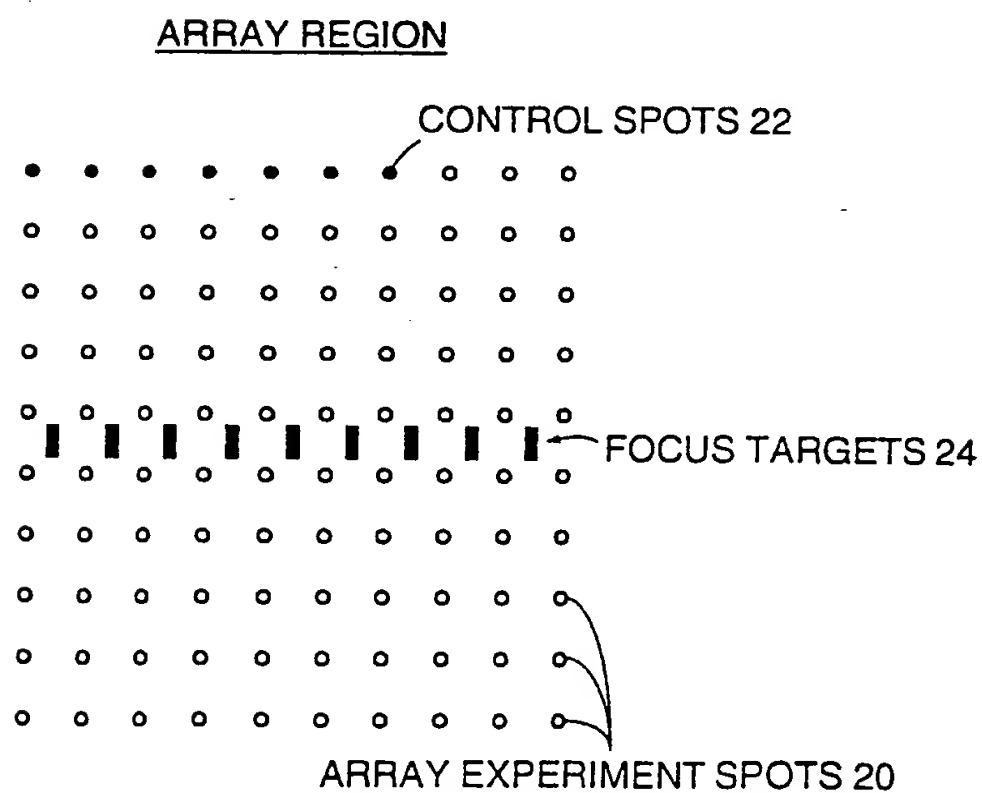


FIG. 2

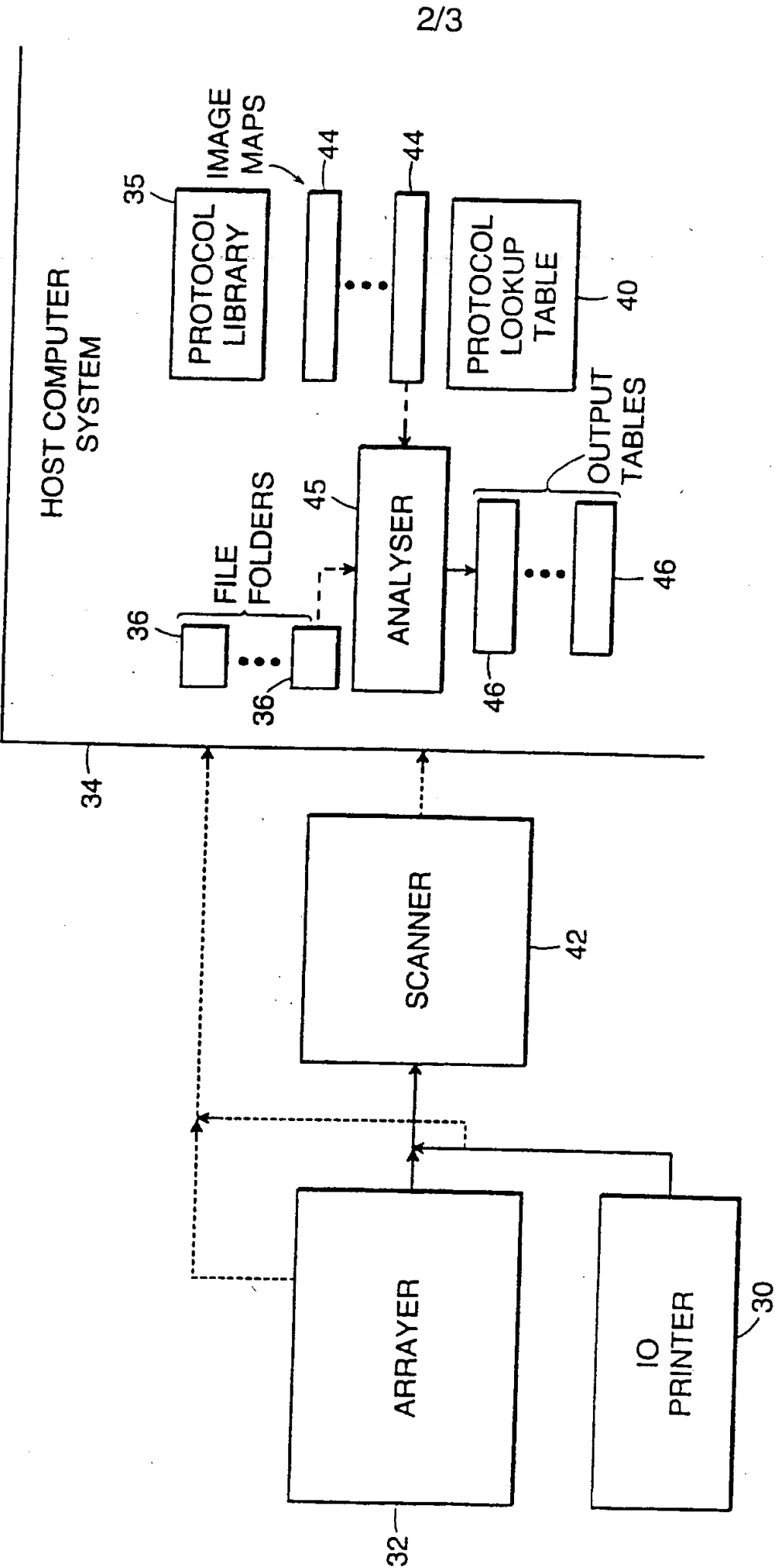
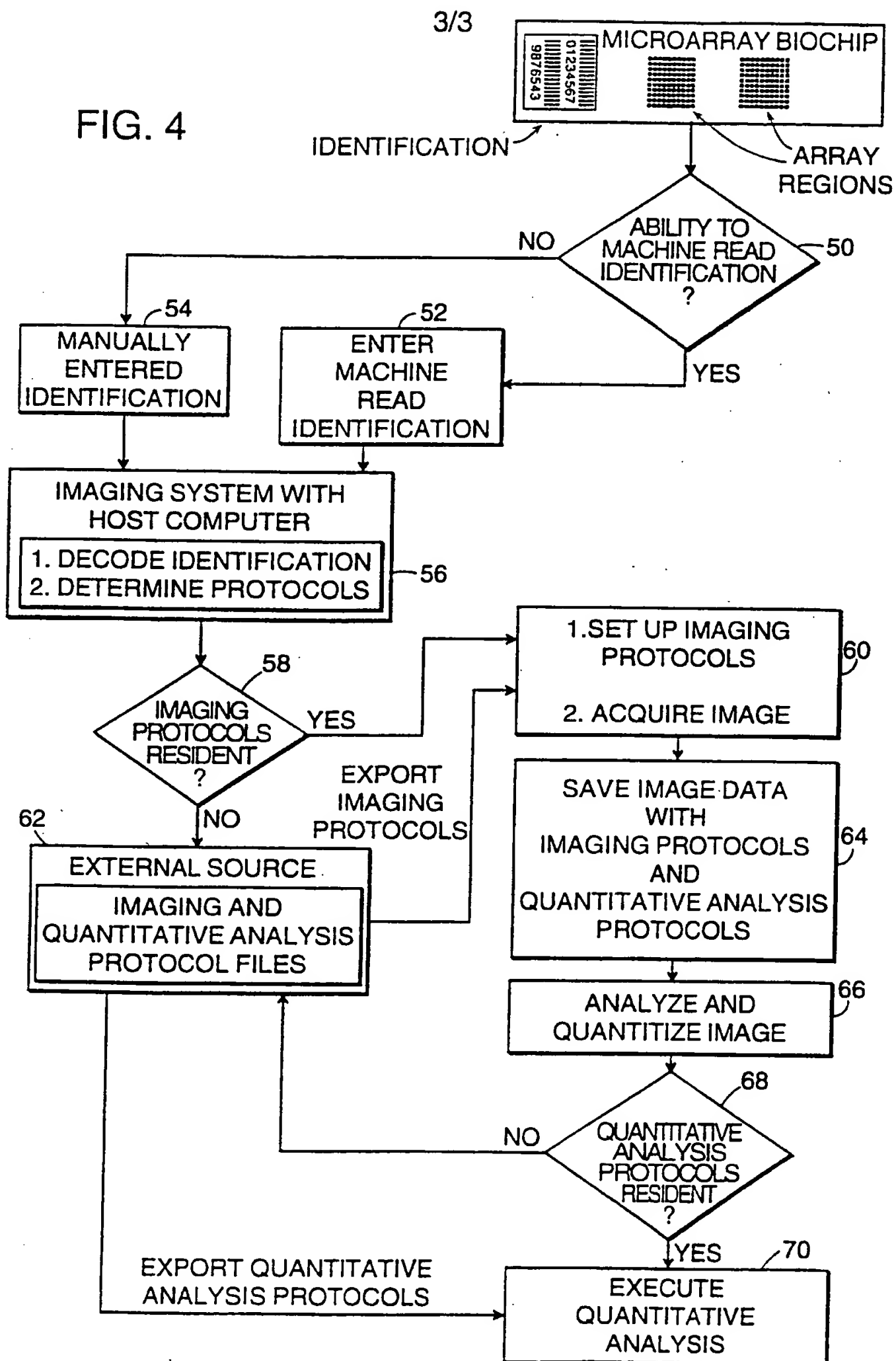


FIG. 3

FIG. 4



# INTERNATIONAL SEARCH REPORT

Inter. Application No  
PCT/US 00/04926

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06F19/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 05591 A (AFFYMETRIX INC) 4 February 1999 (1999-02-04) abstract; claims 1,3 page 9, line 31 -page 10, line 10 page 18, line 4 - line 24	1-4
E	WO 00 23803 A (MILLSTEIN LARRY S) 27 April 2000 (2000-04-27) page 11, line 18 - line 31; figures 6,8A page 53, line 6 -page 54, line 27	1-4
A	US 5 571 639 A (HUBBELL EARL A ET AL) 5 November 1996 (1996-11-05) column 13, line 39 -column 14, line 42	1-4

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

7 August 2000

Date of mailing of the international search report

11/08/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Filloy García, E

# INTERNATIONAL SEARCH REPORT

information on patent family members

Inter xnal Application No

PCT/US 00/04926

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9905591 A	04-02-1999	EP 1002264 A	24-05-2000
		EP 1009861 A	21-06-2000
		EP 0998697 A	10-05-2000
		EP 1007737 A	14-06-2000
		WO 9905323 A	04-02-1999
		WO 9905574 A	04-02-1999
		WO 9905324 A	04-02-1999
WO 0023803 A	27-04-2000	AU 6515899 A	08-05-2000
		AU 9806298 A	03-05-1999
US 5571639 A	05-11-1996	US 5593839 A	14-01-1997
		US 5856101 A	05-01-1999

# INTERNATIONAL SEARCH REPORT

information on patent family members

Inter xnal Application No

PCT/US 00/04926

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9905591 A	04-02-1999	EP 1002264 A	24-05-2000
		EP 1009861 A	21-06-2000
		EP 0998697 A	10-05-2000
		EP 1007737 A	14-06-2000
		WO 9905323 A	04-02-1999
		WO 9905574 A	04-02-1999
		WO 9905324 A	04-02-1999
WO 0023803 A	27-04-2000	AU 6515899 A	08-05-2000
		AU 9806298 A	03-05-1999
US 5571639 A	05-11-1996	US 5593839 A	14-01-1997
		US 5856101 A	05-01-1999